

Deposit

Cosmids 21A4-2-1, 21A4-4-3-1, 21A4-P5-1 have been deposited with the American Type Culture Collection on July 8, 1996, and bear the accession numbers ATCC No. 97649, 97650, and 97651. Plasmid pKExNPR1 was deposited on July 31, 1996 and bears the accession number ATCC No. 97671. Applicants acknowledge their responsibility to replace these plasmids should it loose viability before the end of the term of a patent issued hereon, and their responsibility to notify the American Type Culture Collection of the issuance of such a patent, at which time the deposit will be made available to the public. Prior to that time the deposit will be made available to the Commissioner of Patents under terms of 37 CFR § 1.14 and 35 USC § 112. These deposits are available as required by foreign patent laws in countries wherein counterparts of this subject application, or progeny, are filed. It should be understood that availability of a deposit does not constitute a license to practice the subject invention.

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each independent publication or patent application was specifically and individually indicated to be incorporated by reference.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Dong et al.

(ii) TITLE OF THE INVENTION:
ACQUIRED RESISTANCE GENES AND USES THEREOF

(iii) NUMBER OF SEQUENCES: 28

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Clark & Elbing LLP
(B) STREET: 176 Federal Street
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02110

5 (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

10 (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 60/023,851
(B) FILING DATE: August 9, 1996

(A) APPLICATION NUMBER: 60/035,166
(B) FILING DATE: January 10, 1997

15 (A) APPLICATION NUMBER: 60/046,769
(B) FILING DATE: May 16, 1997

20 (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Elbing, Karen L
(B) REGISTRATION NUMBER: 35,238
(C) REFERENCE/DOCKET NUMBER: 00786/339004

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 617-428-0200
(B) TELEFAX: 617-428-7045

25 (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7548 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

35 AAGCTTGTGA TGCAAGTCAT GGGATATTGC TTTGTGTTAA GTATACAAAA CCATCACGTG 60
GATACATAGT CTTCAAACCA ACCACTAAAC AGTATCAGGT CATACCAAAG CCAGAAGTGA 120
AGGGTTGGGA TATGTCATTG GGTTAGCGG TAATCGGATT GAACCCCTTC CCGTATAAAA 180
TACAAAGGCT TTCGCAGTCT CGGCGTATGT GTATGTCTCG GGGTATCTAC CATTTGAATC 240
ACAGAACTTT TATGTGCGAA GTTTTCGATT CTGATTCGTT TACCTGGAAG AGATTAGAAA 300
TTTGCGTCTA CCAAAAACAG ACAGATTAAT TTTTCCAAAC CCGATACAAG TTTCGGGGTT 360
CTTGCAATTG ATATCACGGA ACAACAATGT GATCCGGTTT TGTCTCAAAA CCGAAACTTG 420
40 GTCCTTCTTC CATACTCCGA ACTCTGATGT TTTCTCAGGA TTAGTCAGAT ACGAAGGGAA 480
GCTAGGTGCT ATTCGTCAGT GGACAAACAA AGATCAAGAA GATGTTTCACG AGTTATGGGT 540

5	TTAAAGAGC	AGTTTTGAAA	AGTCGTGGGT	TAAAGTGAAA	GATATTA AAA	GCATTGGAGT	600
	AGATTTGATT	ACGTGGACTC	CAAGCAACGA	CGTTGTATTG	TTTCGTAGTA	GTGATCGTGG	660
	TTGCCTCTAC	AACATAAACG	CAGAGAAAGT	GAATTTAGTT	TATGCAAAAA	AAGAGGGATC	720
	TGATTGTTCT	TTCGTTTGT	TTCCGTTTTG	TTCTGATTAC	GAGAGGGTTG	ATCTGAACGG	780
	AAGAAGCAAC	GGGCCGACAC	TTTAAAAAAA	AAATAAAAAA	AATGGGCCGA	CAAAATGCAAA	840
10	CGTAGTTGAC	AAGGATCTCA	AGTCTCAAGT	CTCAATTGGC	TCGCTCATTG	TGGGGCATAA	900
	ATATATCTAG	TGATGTTTAA	TTGTTTTTTA	TAAAGTAAAA	AGGAATATTG	AATTTTGTTT	960
	CTTAGTTTA	TGTAATAATA	CGAAACATTG	TTTTATGAAT	ATTTAACTGT	ATTTTTTGGC	1020
	TAGTTATTTT	ATTATATCAA	GGGTTCTGT	TTATAGTTGA	AAACAGTTAC	TGTTATAGAAA	1080
	ATAGTGTCCC	AATTTTCTCT	CTTAAATAAT	ATATTAGTTA	ATAAAAGATA	TTTTAATATA	1140
15	TTAGATATAC	AATAATATCT	AAAGCAACAC	ATATTTAGAC	ACAACACGTA	ATATCTTACT	1200
	ATTGTTTACA	TATATTTATA	GCTTACCAAT	ATAACCCGTA	TCTATGTTTT	ATAAGCTTTT	1260
	ATACAATATA	TGTACGGTAT	GCTGTCCACG	TATATATATT	CTCCAAAAAA	AACGCATGGT	1320
	ACACAAAATT	TATTAATAAT	TTGGCAATTG	GGTGTTTATC	TAAAGTTTAT	CACAAATATT	1380
	ATCAACTATA	ATAGATGGTA	GAAGATAAAA	AAATTATATC	AGATTGATTC	AATTAATAAT	1440
20	TATAATATAT	CATTTTAAAA	AATTAATTA	AAGAAAATA	TTTCATAAAA	TTGTTCAAAA	1500
	GATAATTAGT	AAAATTAAAT	AAATATGTGA	TGCTATGTAA	TTATAGAGAG	TTATTGTAAA	1560
	TTTACTTAAA	ATCATACAAA	TCTTATCCTA	TTTAACTTA	TCATTTAAGA	AATACAAAAG	1620
	TAAAAAACGC	GGAAAGCAAT	AATTTATTTA	CCTTATTATA	ACTCCTATAT	AAAGTACTCT	1680
	GTTTATTCAA	CATAATCTTA	CGTTGTTGTA	TTCATAGGCA	TCTTTAACCT	ATCTTTTCAT	1740
25	TTTCTGATCT	CGATCGTTTT	CGATCCAACA	AAATGAGTCT	ACCGGTGAGG	AACCAAGAGG	1800
	TGATTATGCA	GATTCCTTCT	TCTTCTCAGT	TTCCAGCAAC	ATCGAGTCCG	GAAAACACCA	1860
	ATCAAGTGAA	GGATGAGCCA	AATTTGTTTA	GACGTGTTAT	GAATTTGCTT	TTACGTCGTA	1920
	GTTATTGAAA	AAGCTGATTT	ATCGCATGAT	TCAGAACGAG	AAGTTGAAGG	CAAAATAACTA	1980
	AAGAAGTCTT	TTATATGTAT	ACAATAATTG	TTTTTAAATC	AAATCCTAAT	TAAAAAATA	2040
30	TATTCAATTAT	CGTTTTCATG	TTTTTAATGT	AATTTATTC	TATATCTATA	ATGATTTTGT	2100
	TTGTGAAGAG	GACTTTTCATT	TGCTATAGAA	CAAGGAGAA	AGTTCCAGTA	AATATTGAC	2160
	TTGATTTAAT	TATAGTGTA	AACATGCTGAA	CACGTGAAAT	TACTTTTCA	ATAAACGAAA	2220
	AATATAATAT	ACATTACAAA	ACTTATGTGA	ATAAAGCATG	AGACTTAATA	TACGTTCCCT	2280
	TTATCATTTT	ACTTCAAAGA	AAATAAACAG	AAATGTAAC	TTCACATGTA	AATCTAATTC	2340
35	TTAAATTTAA	AAAATAATAT	TTATATATTT	ATATGAAAAT	AACGAACCGG	ATGAAAAATA	2400
	AATTTTATAT	ATTTTATATCA	TCTCCAAATC	TAGTTTGGTT	CAGGGGCTTA	CCGAACCGGA	2460
	TTGAAC TTCT	CATATACAAA	AATTAGCAAC	ACAAAATGTC	TCCGGTATAA	ATACTAACAT	2520
	TTATAACCCG	AACCGGTTTA	GCTTCCTGTT	ATATCTTTT	AAAAAGATC	TCTGACAAAG	2580
	ATTCCTTTCT	TGGAQAATTA	CCGGTTTGGG	TGAAATGTAA	ACCGTGGGAC	GAGGATGCTT	2640
40	CTTCATATCT	CACCACCACT	CTCGTTGACT	GGACTTGGCT	CTGCTCGTCA	ATGTTTATCT	2700
	TCGATCTTAA	ACCAAATCCA	GTTGATAAGG	TCTCTTCGTT	GATTAGCAGA	GATCTCTTTA	2760
	ATTTGTGAAT	TTCAATTCAT	CGGAACCTGT	TGATGGACAC	CACCATTGAT	GGATTGCGCG	2820
	ATTCTTATGA	AATCAGCAGC	ACTAGTTTCG	TCGCTACCGA	TAACACCGAC	TCCTCTATTG	2880
	TTTATCTGGC	CGCCGAACAA	GTACTCACCG	GACCTGATGT	ATCTGCTCTG	CAATTGCTCT	2940
45	CCAACAGCTT	CGAATCCGTC	TTTGACTCGC	CGGATGATTT	CTACAGCGAC	GCTAAGCTTG	3000
	TTCTCTCCGA	CGGCCGGGAA	GTTTCTTTCC	ACCGGTGCGT	TTTGTCAGCG	AGAAGCTCTT	3060
	TCTTCAAGAG	CGCTTTAGCC	GCCGCTAAGA	AGGAGAAAGA	CTCCAACAAC	ACCGCCGCCG	3120
	TGAAGCTCGA	GCTTAAGGAG	ATTGCCAAGG	ATTACGAAGT	CGGTTTCGAT	TCGGTTGTGA	3180
	CTGTTTTTGGC	TTATGTTTAC	ACAGCAGAG	TGAGAACCGC	GCCTAAAGGA	GTTTCTGAAT	3240
50	GCGCAGACGA	GAATTGCTGC	CACGTGGCTT	GCCGGCCGGC	GGTGATTTC	ATGTTGGAGG	3300
	TTCTCTATTT	GGCTTTCATC	TTCAAGATCC	CTGAATTAAT	TACTCTCTAT	CAGGTAAAAC	3360
	ACCATCTGCA	TTAAGCTATG	GTTACACATT	CATGAATATG	TTCTTACTTG	AGTACTTGTA	3420
	TTTGTATTTT	AGAGGCACTT	ATTGGACGTT	GTAGACAAAG	TTGTTATAGA	GGACACATTG	3480
	GTTTACTCTA	AGCTTGCTAA	TATATGTGGT	AAAGCTTGTA	TGAAGCTATT	GGATAGATGT	3540
55	AAAGAGATTA	TTGTCAAGTC	TAATGTAGAT	ATGGTTAGTC	TTGAAAAGTC	ATTGCCGGAA	3600
	GAGCTTGTTA	AAGAGATAAT	TGATAGACGT	AAAGAGCTTG	GTTTGGAGGT	ACCTAAAGTA	3660
	AAGAACAATG	TCTCGAATGT	ACATAAGGCA	CTTGACTCGG	ATGATATTGA	GTTAGTCAAG	3720
	TTGCTTTTGA	AAGAGGATCA	CACCAATCTA	GATGATGCTG	GTGCTTTTCA	TTTCGCTGTT	3780
	GCATATTGCA	ATGTGAAGAC	CGCAACAGAT	CTTTTAAAC	TTGATCTTGC	CGATGTCAAC	3840

	CATAGGAATC	CGAGGGGATA	TACGGTGCTT	CATGTTGCTG	CGATGCGGAA	GGAGCCACAA	3900
	TTGATACTAT	CTCTATTGGA	AAAAGGTGCA	AGTGCATCAG	AAGCAACTTT	GGAAGGTAGA	3960
	ACCGCACTCA	TGATCGCAAA	ACAAGCCACT	ATGGCGGTTG	AATGTAATAA	TATCCCGGAG	4020
5	CAATGCAAGC	ATTCTCTCAA	AGGCCGACTA	TGTGTAGAAA	TACTAGAGCA	AGAAGACAAA	4080
	CGAGAACAAA	TTCCTAGAGA	TGTTCTCTCC	TCTTTTGCAG	TGGCGGCCGA	TGAATTGAAG	4140
	ATGACGCTGC	TCGATCTTGA	AAATAGAGGT	ATCTATCAAG	TCTTATTTCT	TATATGTTTG	4200
	AATTAAATTT	ATGTCCTCTC	TATTAGGAAA	CTGAGTGAAC	TAATGATAAC	TATTCTTTGT	4260
	GTCGTCCACT	GTTTAGTTGC	ACTTGCTCAA	CGTCTTTTTC	CAACGGAAGC	ACAAGCTGCA	4320
	ATGGAGATCG	CCGAAATGAA	GGGAACATGT	GAGTTCATAG	TGACTAGCCT	CGAGCCTGAC	4380
10	CGTCTCACTG	GTACGAAGAG	AACATCACCG	GGTGTAAGA	TAGCACCTTT	CAGAATCCTA	4440
	GAAGAGCATC	AAAGTAGACT	AAAAGCGCTT	TCTAAAACCG	GTATGGATTC	TCACCCACTT	4500
	CATCGGACTC	CTTATCACAA	AAAACAAAAC	TAAATGATCT	TTAAACATGG	TTTTGTACT	4560
	TGCTGTCTGA	CTTGTGTTTT	TTATCATCAG	TGGAACCTCG	GAAACGATTC	TTCCCGCGCT	4620
	GTTCCGCAGT	GCTCGACCAG	ATTATGAACT	GTGAGGACTT	GACTCAACTG	GCTTGCGGAG	4680
15	AAGACGACAC	TGCTGAAGAA	ACGACTACAA	AAGAAGCAAA	GGTACATGGA	AATACAAGAG	4740
	ACACTAAAGA	AGGCCTTTAG	TGAGGACAAT	TTGGAATTAG	GAAATTCGTC	CCTGACAGAT	4800
	TCGACTTCTT	CCACATCGAA	ATCAACCGGT	GGAAAAGAGT	CTAACCCTAA	ACTCTCTCAT	4860
	CGTCGTCGGT	GAGACTCTTG	CCTCTTAGTG	TAATTTTTGC	TGTACCATAT	AATTCTGTTT	4920
	TCATGATGAC	TGTAAGTGT	TATGTCTATC	GTTGGCGTCA	TATAGTTTCG	CTCTTCGTTT	4980
20	TGCATCTGT	GTATTATTGC	TGCAGGTGTG	CTTCAAACAA	ATGTTGTAAC	AATTTGAACC	5040
	AATGATATAC	AGATTGTGTA	TATATATTTA	TGTACATCAA	CAATAACCCA	TGATGGTGT	5100
	ACAGAGTTGC	TAGAATCAAA	GTGTGAAATA	ATGTCAAATT	GTTTCTCTGT	TGGATATTTT	5160
	CCACCAAGAA	CCAAAAGAAT	ATTCAAGTTC	CCTGAACCTC	TGGCAACATT	CATGTTATAT	5220
	GTATCTTCTT	AATTCTTCTT	TTAACCTTTT	GTAACCTCGA	TTACACAGCA	AGTTAGTTTC	5280
25	AGGTCTAGAG	ATAAGAGAAC	ACTGAGTGGG	CGTGTAAGGT	GCATTCTCCT	AGTCAGCTCC	5340
	ATTGCATCCA	ACATTGTGTA	ATGACACAAG	TTAACAATCC	TTTGCACCAT	TTCTGGGTGC	5400
	ATACATGGAA	ACTTCTTCGA	TTGAAACTTC	CCACATGTGC	AGGTGCGTTC	GCTGTCACTG	5460
	ATAGACCAAG	AGACTGAAAG	CTTTCACAAA	TTGCCCTCAA	ATCTTCTGTT	TCTATCGTCA	5520
	TGACTCCATA	TCTCCGACCA	CTGGTCAATG	GCCAGAGCCC	ACTGATTTTG	AGGGAATTGG	5580
30	GCTAACCATT	TCCGAGCTTC	TGAGTCTTTC	TTTTTGATGT	CCTTTATGTA	GGAATCAAAT	5640
	TCTTCTTCT	GACTTGTGGA	TCCAGCCTGC	TTCAACAAGC	TCACCAGGTT	GTAGTCTCCA	5700
	AAAATATCAT	GGAATTGTAA	GCAAAAACAA	TCCAGACAGA	ACCTGTGATA	GACCCAAGGT	5760
	TCTTGCCACA	GTGATCCGGG	TTCGTTAATA	ACAGCAACTA	TGTCGGGGTG	AGGACTGGAG	5820
	ACGAAGCAAA	CGTCTTCTCT	TTGTGTTACC	TTCTCTCTGA	TATTAGTGAG	AAACCAACGC	5880
35	CAACTATCAG	TGGACACTTC	TTTGGTAAGC	GGAAAGCAAG	CGGGAAAAAC	AATCATCAGC	5940
	GTCGAGTCCT	GAGGAAAATC	ATCAATTTC	TAGGGGTACT	TGCCGTTCAA	GTCTTTTGAA	6000
	TCCACTATGA	TCAGAGGTCT	ACAGTGTGTA	AACCCCTCAA	TGGACTGTGG	AAACGCCCAA	6060
	AACGCGCCAC	CGAAGGATGC	AAATTCAGGA	TTAGGGAAAA	GCTCATATTG	CAGTCCACAA	6120
	GTAGCCCAT	AGATGAGTGA	AATGCAGCCA	ATTAGTTTAG	GCAATACTCT	GAAACTCTGA	6180
40	TCTTTGATTA	CTTCTGTTC	TGCTGCCCGC	AGCTTTGAAG	TTTAAAGCAT	GTCACCAAAC	6240
	TTTTCAACTC	TGCTGTTAGA	GTGGGTGTGA	CCCTGATCAG	ACACTCAATC	TCTTCTGCTG	6300
	CAAATTACAA	GTTGAAGTTT	TCCGGCTTAA	TAGAACAACA	AGTATGTGGA	CCAACTACAC	6360
	TTAGTTATCT	TAACAAGTCC	ATGTTCTTCT	ATTCAATCTG	CCCGACGCGA	CCAATTGCAT	6420
	TTCCATCTGA	TGCATTTAAA	CGTATACTCG	TCTTCTCAA	TCTTGTGAC	TACACACTTT	6480
45	TGCTGCCCTC	TAATGGAACA	CCAGTCCACC	GCCTTCTTCA	GCTCATCCCT	ATCTTTAAAA	6540
	CACAACCCTA	CACGCAATTC	ATGATCATCA	ATCCACAAAC	TAGACAAAGT	ACACTGTTTT	6600
	GAAGCACTCG	AATCAACAAC	ACCTTACTTT	AATAAGCACG	CATACGGTAA	TACCTCTAAG	6660
	CCTGGCACAT	TCAAACCTTG	TGTGCATCAT	CTGAACCCGA	GTTTTTATCC	GTTATTTCTC	6720
	CATCCCCACC	TCCACGAGTG	CTACCATTTT	CGAAGTCAGA	ATTTTCCTCG	TCTTCAATCC	6780
50	ACCCGTTACT	GTTACCCACT	CCCTGAACCT	CTAAACCAT	ATCTCTCTCT	ACTTTCACAG	6840
	ATGCATGTGA	CACATAATCA	GTAGCTTCTT	GGGGTTGTTG	CGTCCTCTGT	GTATTCGAGG	6900
	AACTAGCGGG	ATATTCTATT	ACGGATGAAC	AAGCAGCATG	ATCAGTAACA	TTATCAGATG	6960
	TCGATTTCAC	TTCCAAATAC	AACTCCACAT	TTCTTATAGA	AGGATGATAA	CTTGGAACCT	7020
	CAAGCATAGT	CTCCAAACTA	GTGTGTTTCA	CTACATGAAG	AAGTAGATAG	ATAAAGAGAT	7080
55	CCGGTGAAAC	AACTACAGGA	TACTTACCAA	AATATATTGA	ACACTGATTT	CTGCAGCTGC	7140

5 AATCCAAAAA TTGGATAAAG ACCATTCAAC AATGTACTTA ACGCAGTCTT TTGCCTAACC 7200
 TTGACCGTTT TAGGAGTGGA TCCTTCATAG TAAACACCAT CAGGACCATA CTTGGTAGAA 7260
 CCTTTCTCTC AAGGTTTCCA TCGCCATGAC CATAACAGTC CTGCAGTGAA TTCTAAGAAA 7320
 AATGTAAAAA ATTTTGGCCT AAACCTATAA TTCTTAACAT ACGAAACCAT GGAGAACTCC 7380
 ATGTCTAAAA AATAAAGGCT AAAGCTTTTT GCGCAGAGAA GCAGATAAAT CCATTCAAAA 7440
 CACATAAACT CTAAACAATA AACAGTGATA CTCAATACTA AGACTTGTAAG AGGTCTACGT 7500
 AACTCAAAAC TGGAGAATTG TCAGATCGGG TGTGGCTAGT AGAAGCTT 7548

(2) INFORMATION FOR SEQ ID NO:2:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2104 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: cDNA
 (ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 93...1871
 (D) OTHER INFORMATION:

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TCGATCTTTA ACCAAATCCA GTTGATAAGG TCTCTTCGTT GATTAGCAGA GATCTCTTTA 60
 ATTTGTGAAT TTCAATTCAT CGGAACCTGT TG ATG GAC ACC ACC ATT GAT GGA 113
 Met Asp Thr Thr Ile Asp Gly
 1 5

25 TTC GCC GAT TCT TAT GAA ATC AGC AGC ACT AGT TTC GTC GCT ACC GAT 161
 Phe Ala Asp Ser Tyr Glu Ile Ser Ser Thr Ser Phe Val Ala Thr Asp
 10 15 20

30 AAC ACC GAC TCC TCT ATT GTT TAT CTG GCC GCC GAA CAA GTA CTC ACC 209
 Asn Thr Asp Ser Ser Ile Val Tyr Leu Ala Ala Glu Gln Val Leu Thr
 25 30 35

GGA CCT GAT GTA TCT GCT CTG CAA TTG CTC TCC AAC AGC TTC GAA TCC 257
 Gly Pro Asp Val Ser Ala Leu Gln Leu Leu Ser Asn Ser Phe Glu Ser
 40 45 50 55

35 GTC TTT GAC TCG CCG GAT GAT TTC TAC AGC GAC GCT AAG CTT GTT CTC 305
 Val Phe Asp Ser Pro Asp Asp Phe Tyr Ser Asp Ala Lys Leu Val Leu
 60 65 70

TCC GAC GGC CGG GAA GTT TCT TTC CAC CGG TGC GTT TTG TCA GCG AGA 353
 Ser Asp Gly Arg Glu Val Ser Phe His Arg Cys Val Leu Ser Ala Arg
 75 80 85

40 AGC TCT TTC TTC AAG AGC GCT TTA GCC GCC GCT AAG AAG GAG AAA GAC 401
 Ser Ser Phe Phe Lys Ser Ala Leu Ala Ala Ala Lys Lys Glu Lys Asp
 90 95 100

	TCC AAC AAC ACC GCC GCC GTG AAG CTC GAG CTT AAG GAG ATT GCC AAG	449
	Ser Asn Asn Thr Ala Ala Val Lys Leu Glu Leu Lys Glu Ile Ala Lys	
	105 110 115	
5	GAT TAC GAA GTC GGT TTC GAT TCG GTT GTG ACT GTT TTG GCT TAT GTT	497
	Asp Tyr Glu Val Gly Phe Asp Ser Val Val Thr Val Leu Ala Tyr Val	
	120 125 130 135	
	TAC AGC AGC AGA GTG AGA CCG CCG CCT AAA GGA GTT TCT GAA TGC GCA	545
	Tyr Ser Ser Arg Val Arg Pro Pro Pro Lys Gly Val Ser Glu Cys Ala	
	140 145 150	
10	GAC GAG AAT TGC TGC CAC GTG GCT TGC CGG CCG GCG GTG GAT TTC ATG	593
	Asp Glu Asn Cys Cys His Val Ala Cys Arg Pro Ala Val Asp Phe Met	
	155 160 165	
	TTG GAG GTT CTC TAT TTG GCT TTC ATC TTC AAG ATC CCT GAA TTA ATT	641
	Leu Glu Val Leu Tyr Leu Ala Phe Ile Phe Lys Ile Pro Glu Leu Ile	
15	170 175 180	
	ACT CTC TAT CAG AGG CAC TTA TTG GAC GTT GTA GAC AAA GTT GTT ATA	689
	Thr Leu Tyr Gln Arg His Leu Leu Asp Val Val Asp Lys Val Val Ile	
	185 190 195	
20	GAG GAC ACA TTG GTT ATA CTC AAG CTT GCT AAT ATA TGT GGT AAA GCT	737
	Glu Asp Thr Leu Val Ile Leu Lys Leu Ala Asn Ile Cys Gly Lys Ala	
	200 205 210 215	
	TGT ATG AAG CTA TTG GAT AGA TGT AAA GAG ATT ATT GTC AAG TCT AAT	785
	Cys Met Lys Leu Leu Asp Arg Cys Lys Glu Ile Ile Val Lys Ser Asn	
	220 225 230	
25	GTA GAT ATG GTT AGT CTT GAA AAG TCA TTG CCG GAA GAG CTT GTT AAA	833
	Val Asp Met Val Ser Leu Glu Lys Ser Leu Pro Glu Glu Leu Val Lys	
	235 240 245	
	GAG ATA ATT GAT AGA CGT AAA GAG CTT GGT TTG GAG GTA CCT AAA GTA	881
	Glu Ile Ile Asp Arg Arg Lys Glu Leu Gly Leu Glu Val Pro Lys Val	
30	250 255 260	
	AAG AAA CAT GTC TCG AAT GTA CAT AAG GCA CTT GAC TCG GAT GAT ATT	929
	Lys Lys His Val Ser Asn Val His Lys Ala Leu Asp Ser Asp Asp Ile	
	265 270 275	
35	GAG TTA GTC AAG TTG CTT TTG AAA GAG GAT CAC ACC AAT CTA GAT GAT	977
	Glu Leu Val Lys Leu Leu Leu Lys Glu Asp His Thr Asn Leu Asp Asp	
	280 285 290 295	
	GCG TGT GCT CTT CAT TTC GCT GTT GCA TAT TGC AAT GTG AAG ACC GCA	1025
	Ala Cys Ala Leu His Phe Ala Val Ala Tyr Cys Asn Val Lys Thr Ala	
	300 305 310	
40	ACA GAT CTT TTA AAA CTT GAT CTT GCC GAT GTC AAC CAT AGG AAT CCG	1073
	Thr Asp Leu Leu Lys Leu Asp Leu Ala Asp Val Asn His Arg Asn Pro	
	315 320 325	

	AGG GGA TAT ACG GTG CTT CAT GTT GCT GCG ATG CGG AAG GAG CCA CAA	1121
	Arg Gly Tyr Thr Val Leu His Val Ala Ala Met Arg Lys Glu Pro Gln	
	330 335 340	
5	TTG ATA CTA TCT CTA TTG GAA AAA GGT GCA AGT GCA TCA GAA GCA ACT	1169
	Leu Ile Leu Ser Leu Leu Glu Lys Gly Ala Ser Ala Ser Glu Ala Thr	
	345 350 355	
	TTG GAA GGT AGA ACC GCA CTC ATG ATC GCA AAA CAA GCC ACT ATG GCG	1217
	Leu Glu Gly Arg Thr Ala Leu Met Ile Ala Lys Gln Ala Thr Met Ala	
	360 365 370 375	
10	GTT GAA TGT AAT AAT ATC CCG GAG CAA TGC AAG CAT TCT CTC AAA GGC	1265
	Val Glu Cys Asn Asn Ile Pro Glu Gln Cys Lys His Ser Leu Lys Gly	
	380 385 390	
	CGA CTA TGT GTA GAA ATA CTA GAG CAA GAA GAC AAA CGA GAA CAA ATT	1313
15	Arg Leu Cys Val Glu Ile Leu Glu Gln Glu Asp Lys Arg Glu Gln Ile	
	395 400 405	
	CCT AGA GAT GTT CCT CCC TCT TTT GCA GTG GCG GCC GAT GAA TTG AAG	1361
	Pro Arg Asp Val Pro Pro Ser Phe Ala Val Ala Ala Asp Glu Leu Lys	
	410 415 420	
20	ATG ACG CTG CTC GAT CTT GAA AAT AGA GTT GCA CTT GCT CAA CGT CTT	1409
	Met Thr Leu Leu Asp Leu Glu Asn Arg Val Ala Leu Ala Gln Arg Leu	
	425 430 435	
	TTT CCA ACG GAA GCA CAA GCT GCA ATG GAG ATC GCC GAA ATG AAG GGA	1457
	Phe Pro Thr Glu Ala Gln Ala Ala Met Glu Ile Ala Glu Met Lys Gly	
	440 445 450 455	
25	ACA TGT GAG TTC ATA GTG ACT AGC CTC GAG CCT GAC CGT CTC ACT GGT	1505
	Thr Cys Glu Phe Ile Val Thr Ser Leu Glu Pro Asp Arg Leu Thr Gly	
	460 465 470	
	ACG AAG AGA ACA TCA CCG GGT GTA AAG ATA GCA CCT TTC AGA ATC CTA	1553
30	Thr Lys Arg Thr Ser Pro Gly Val Lys Ile Ala Pro Phe Arg Ile Leu	
	475 480 485	
	GAA GAG CAT CAA AGT AGA CTA AAA GCG CTT TCT AAA ACC GTG GAA CTC	1601
	Glu Glu His Gln Ser Arg Leu Lys Ala Leu Ser Lys Thr Val Glu Leu	
	490 495 500	
35	GGG AAA CGA TTC TTC CCG CGC TGT TCG GCA GTG CTC GAC CAG ATT ATG	1649
	Gly Lys Arg Phe Phe Pro Arg Cys Ser Ala Val Leu Asp Gln Ile Met	
	505 510 515	
	AAC TGT GAG GAC TTG ACT CAA CTG GCT TGC GGA GAA GAC GAC ACT GCT	1697
	Asn Cys Glu Asp Leu Thr Gln Leu Ala Cys Gly Glu Asp Asp Thr Ala	
	520 525 530 535	
40	GAG AAA CGA CTA CAA AAG AAG CAA AGG TAC ATG GAA ATA CAA GAG ACA	1745
	Glu Lys Arg Leu Gln Lys Lys Gln Arg Tyr Met Glu Ile Gln Glu Thr	
	540 545 550	

CTA AAG AAG GCC TTT AGT GAG GAC AAT TTG GAA TTA GGA AAT TCG TCC 1793
 Leu Lys Lys Ala Phe Ser Glu Asp Asn Leu Glu Leu Gly Asn Ser Ser
 555 560 565

5 CTG ACA GAT TCG ACT TCT TCC ACA TCG AAA TCA ACC GGT GGA AAG AGG 1841
 Leu Thr Asp Ser Thr Ser Ser Thr Ser Lys Ser Thr Gly Gly Lys Arg
 570 575 580

TCT AAC CGT AAA CTC TCT CAT CGT CGT CGG TGAGACTCTT GCCTCTTAGT GTA 1894
 Ser Asn Arg Lys Leu Ser His Arg Arg Arg
 585 590

10 ATTTTGTCTG TACCATATAA TTCTGTTTTT ATGATGACTG TAACTGTTTA TGTCTATCGT 1954
 TGGCGTCATA TAGTTTCGCT CTTTCGTTTTG CATCCTGTGT ATTATTGCTG CAGGTGTGCT 2014
 TCAAACAAAT GTTGTAAACAA TTTGAACCAA TGGTATACAG ATTTGTAATA TATATTTATG 2074
 TACATCAACA ATAAAAAAAA AAAAAAAAAA 2104

(2) INFORMATION FOR SEQ ID NO:3:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 593 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Asp Thr Thr Ile Asp Gly Phe Ala Asp Ser Tyr Glu Ile Ser Ser
 1 5 10 15
 25 Thr Ser Phe Val Ala Thr Asp Asn Thr Asp Ser Ser Ile Val Tyr Leu
 20 25 30
 Ala Ala Glu Gln Val Leu Thr Gly Pro Asp Val Ser Ala Leu Gln Leu
 35 40 45
 Leu Ser Asn Ser Phe Glu Ser Val Phe Asp Ser Pro Asp Asp Phe Tyr
 50 55 60
 30 Ser Asp Ala Lys Leu Val Leu Ser Asp Gly Arg Glu Val Ser Phe His
 65 70 75 80
 Arg Cys Val Leu Ser Ala Arg Ser Ser Phe Phe Lys Ser Ala Leu Ala
 85 90 95
 35 Ala Ala Lys Lys Glu Lys Asp Ser Asn Asn Thr Ala Ala Val Lys Leu
 100 105 110
 Glu Leu Lys Glu Ile Ala Lys Asp Tyr Glu Val Gly Phe Asp Ser Val
 115 120 125
 Val Thr Val Leu Ala Tyr Val Tyr Ser Ser Arg Val Arg Pro Pro Pro
 130 135 140
 40 Lys Gly Val Ser Glu Cys Ala Asp Glu Asn Cys Cys His Val Ala Cys
 145 150 155 160
 Arg Pro Ala Val Asp Phe Met Leu Glu Val Leu Tyr Leu Ala Phe Ile
 165 170 175
 45 Phe Lys Ile Pro Glu Leu Ile Thr Leu Tyr Gln Arg His Leu Leu Asp
 180 185 190
 Val Val Asp Lys Val Val Ile Glu Asp Thr Leu Val Ile Leu Lys Leu
 195 200 205

	Ala	Asn	Ile	Cys	Gly	Lys	Ala	Cys	Met	Lys	Leu	Leu	Asp	Arg	Cys	Lys
	210						215					220				
	Glu	Ile	Ile	Val	Lys	Ser	Asn	Val	Asp	Met	Val	Ser	Leu	Glu	Lys	Ser
	225					230					235					240
5	Leu	Pro	Glu	Glu	Leu	Val	Lys	Glu	Ile	Ile	Asp	Arg	Arg	Lys	Glu	Leu
					245						250				255	
	Gly	Leu	Glu	Val	Pro	Lys	Val	Lys	Lys	His	Val	Ser	Asn	Val	His	Lys
				260						265				270		
10	Ala	Leu	Asp	Ser	Asp	Asp	Ile	Glu	Leu	Val	Lys	Leu	Leu	Leu	Lys	Glu
			275					280					285			
	Asp	His	Thr	Asn	Leu	Asp	Asp	Ala	Cys	Ala	Leu	His	Phe	Ala	Val	Ala
	290						295					300				
	Tyr	Cys	Asn	Val	Lys	Thr	Ala	Thr	Asp	Leu	Leu	Lys	Leu	Asp	Leu	Ala
	305					310					315					320
15	Asp	Val	Asn	His	Arg	Asn	Pro	Arg	Gly	Tyr	Thr	Val	Leu	His	Val	Ala
					325						330				335	
	Ala	Met	Arg	Lys	Glu	Pro	Gln	Leu	Ile	Leu	Ser	Leu	Leu	Glu	Lys	Gly
				340					345					350		
20	Ala	Ser	Ala	Ser	Glu	Ala	Thr	Leu	Glu	Gly	Arg	Thr	Ala	Leu	Met	Ile
			355					360					365			
	Ala	Lys	Gln	Ala	Thr	Met	Ala	Val	Glu	Cys	Asn	Asn	Ile	Pro	Glu	Gln
	370						375					380				
	Cys	Lys	His	Ser	Leu	Lys	Gly	Arg	Leu	Cys	Val	Glu	Ile	Leu	Glu	Gln
	385					390					395					400
25	Glu	Asp	Lys	Arg	Glu	Gln	Ile	Pro	Arg	Asp	Val	Pro	Pro	Ser	Phe	Ala
					405					410					415	
	Val	Ala	Ala	Asp	Glu	Leu	Lys	Met	Thr	Leu	Leu	Asp	Leu	Glu	Asn	Arg
				420					425					430		
30	Val	Ala	Leu	Ala	Gln	Arg	Leu	Phe	Pro	Thr	Glu	Ala	Gln	Ala	Ala	Met
			435					440					445			
	Glu	Ile	Ala	Glu	Met	Lys	Gly	Thr	Cys	Glu	Phe	Ile	Val	Thr	Ser	Leu
	450						455					460				
	Glu	Pro	Asp	Arg	Leu	Thr	Gly	Thr	Lys	Arg	Thr	Ser	Pro	Gly	Val	Lys
	465					470					475					480
35	Ile	Ala	Pro	Phe	Arg	Ile	Leu	Glu	Glu	His	Gln	Ser	Arg	Leu	Lys	Ala
					485					490					495	
	Leu	Ser	Lys	Thr	Val	Glu	Leu	Gly	Lys	Arg	Phe	Phe	Pro	Arg	Cys	Ser
				500					505					510		
40	Ala	Val	Leu	Asp	Gln	Ile	Met	Asn	Cys	Glu	Asp	Leu	Thr	Gln	Leu	Ala
			515					520					525			
	Cys	Gly	Glu	Asp	Asp	Thr	Ala	Glu	Lys	Arg	Leu	Gln	Lys	Lys	Gln	Arg
	530						535					540				
	Tyr	Met	Glu	Ile	Gln	Glu	Thr	Leu	Lys	Lys	Ala	Phe	Ser	Glu	Asp	Asn
	545					550					555					560
45	Leu	Glu	Leu	Gly	Asn	Ser	Ser	Leu	Thr	Asp	Ser	Thr	Ser	Ser	Thr	Ser
					565					570					575	
	Lys	Ser	Thr	Gly	Gly	Lys	Arg	Ser	Asn	Arg	Lys	Leu	Ser	His	Arg	Arg
				580					585					590		
50	Arg															

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 49 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Asn His Arg Asn Pro Arg Gly Tyr Thr Val Leu His Val Ala Ala Met
1 5 10 15
Arg Lys Glu Pro Gln Leu Ile Leu Ser Leu Leu Glu Lys Gly Ala Ser
20 25 30
10 Ala Ser Glu Ala Thr Leu Glu Gly Arg Thr Ala Leu Met Ile Ala Lys
35 40 45
Gln

(2) INFORMATION FOR SEQ ID NO:5:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Asn Ala Lys Thr Lys Asn Gly Tyr Thr Ala Leu His Gln Ala Ala Gln
1 5 10 15
25 Gln Gly His Thr His Ile Ile Asn Val Leu Leu Gln Asn Asn Ala Ser
20 25 30
Pro Asn Glu Leu Thr Val Asn Gly Asn Thr Ala Leu Ala Ile Ala Arg
35 40 45
Arg

30 (2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Lys Val Lys Lys His Val Ser Asn Val His Lys Ala Leu Asp Ser Asp
1 5 10 15
40 Asp Ile Glu Leu Val Lys Leu Leu Lys Glu Asp
20 25

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Lys Thr Lys Asn Gly Leu Ser Pro Leu His Met Ala Thr Gln Gly Asp
1 5 10 15
His Leu Asn Cys Val Gln Leu Leu Leu Ser Arg Asn
20 25

10

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Lys His Val Ser Asn Val His Lys Ala Leu Asp Ser Asp Asp Ile Glu
1 5 10 15
Leu Val Lys Leu Leu Leu Lys Glu Asp His Thr Asn Leu Asp Asp Ala
20 25 30
Cys

20

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Asp Asp Ala Cys Ala Leu His Phe Ala Val Ala Tyr Cys Asn Val Lys
1 5 10 15
Thr Ala Thr Asp Leu Leu Lys Leu Asp Leu Ala Asp Val Asn His Arg
20 25 30
Asn

35

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 amino acids
(B) TYPE: amino acid

40

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

5 Arg Gly Tyr Thr Val Leu His Val Ala Ala Met Arg Lys Glu Pro Gln
1 5 10 15
Leu Ile Leu Ser Leu Leu Glu Lys Gly Ala Ser Ala Ser Glu Ala Thr
20 25 30
Leu

10 (2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

20 Glu Gly Arg Thr Ala Leu Met Ile Ala Lys Gln Ala Thr Met Ala Val
1 5 10 15
Glu Cys Asn Asn Ile Pro Glu Gln Cys Lys His Ser Leu Lys Gly Arg
20 25 30
Leu

(2) INFORMATION FOR SEQ ID NO:12:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

30 Gly Thr Pro Leu His Leu Ala Ala Arg Gly His Val Glu Val Val Lys
1 5 10 15
Leu Leu Leu Asp Gly Ala Asp Val Asn Ala Thr Lys Ala Ile Ser Gln
20 25 30
Asn Asn Leu Asp Ile Ala Glu Val Lys Asn Pro Asp Asp Val Lys Thr
35 35 40 45
Met Arg Gln Ser Ile Asn Glu
50 55

(2) INFORMATION FOR SEQ ID NO:13:

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2172 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

	GTGACTTTCT	AACATATGGCT	GAAATTGCAG	AACGAAAAAG	ACTTTCCATT	TTTCACTTGA	60
	ATGAAACCCA	AAATGGAAAT	CTATCTCTCT	TCTTCTTCTC	TTTTACTACC	TCCATTTCCA	120
	TGGCTTTCCC	TCCTCTACCT	TCCCTAGCTC	TTTTCAATTT	CTAGAATAT	CTTTTCTTAG	180
	TCTGTAATTA	TCTATAGCTC	AATTTCTAAG	ACAGAACTTA	TGTAAGGCGG	CTTTCTGTAA	240
10	TGGATAATAG	TAGGACTGCG	TTTCTTGATT	CGAATGACAT	CAGCGGAAGC	AGTAGTATAT	300
	GCTGCATCGG	CGGCGGCATG	ACTGAATTTT	TCTCGCCGGA	GACTTCGCCG	GCGGAGATCA	360
	CTTCACTGAA	ACGCCTATCG	GAAACACTGG	AATCTATCTT	CGATGCGTCT	TTGCCGGAGT	420
	TTGACTACTT	CGCCGACGCT	AAGCTTGTGG	TTTCCGGCCC	GTGTAAGGAA	ATTCCGGTGC	480
	ACCGGTGCAT	TTTGTGCGCG	AGGAGTCCGT	TCTTTAAGAA	TTTGTCTGCG	GGTAAAAAGG	540
15	AGAAGAATAG	TAGTAAGGTG	GAATTGAAGG	AGGTGATGAA	AGAGCATGAG	GTGAGCTATG	600
	ATGCTGTAAT	GAGTGTATTG	GCTTATTTGT	ATAGTGGTAA	AGTTAGGCCT	TCACCTAAAG	660
	ATGTGTGTGT	TTGTGTGGAC	AATGACTGCT	CTCATGTGGC	TTGTAGGCCA	GCTGTGGCAT	720
	TCCTGGTTGA	GGTTTGTGAC	ACATCATTTA	CCTTTCAGAT	CTCTGAATTG	GTTGACAAGT	780
	TTCAGAGACA	CCTACTGGAT	ATTCTTGACA	AAACTGCAGC	AGACGATGTA	ATGATGGTTT	840
20	TATCTGTTGC	AAACATTTGT	GGTAAAGCAT	GCGAGAGATT	GCTTTCAAGC	TGCATTGAGA	900
	TTATTGTCAA	GTCTAATGTT	GATATCATAA	CCCTTGATAA	AGCCTTGCCT	CATGACATTG	960
	TAAACAAAT	TACTGATTCA	CGAGCGGAAC	TTGGTCTACA	AGGGCCTGAA	AGCAACGGTT	1020
	TTCTTGATAA	ACATGTTAAG	AGGATACATA	GGGCATTGGA	TTCTGATGAT	GTTGAATTAC	1080
	TACAAATGTT	GCTAAGAGAG	GGGCATACTA	CCCTAGATGA	TGCATATGCT	CTCCATTATG	1140
25	CTGTAGCGTA	TTGCGATGCA	AAGACTACAG	CAGAACTTCT	AGATCTTGCA	CTTGCTGATA	1200
	TTAATCATCA	AAATTCAAGG	GGATACACGG	TGCTGCATGT	TGCAGCCATG	AGGAAAGAGC	1260
	CTAAAATTGT	AGTGTCCCTT	TTAACCAAAG	GAGCTAGACC	TTCTGATCTG	ACATCCGATG	1320
	GAAGAAAAGC	ACTTCAAATC	GCCAAAGAGG	TCACTAGGCT	TGTGGATTTT	AGTAAGTCTC	1380
	CGGAGGAAGG	AAAATCTGCT	TCGAATGATC	GGTTATGCAT	TGAGATTCTG	GAGCAAGCAG	1440
30	AAAGAAGAGA	CCCTCTGCTA	GGAGAAGCTT	CTGTATCTCT	TGCTATGGCA	GGCGATGATT	1500
	TGCGTATGAA	GCTGTTATAC	CTTGAAAATA	GAGTTGGCCT	GGCTAAACTC	CTTTTCCCAA	1560
	TGGAAGCTAA	AGTTGCAATG	GACATTGCTC	AAGTTGATGG	CACTTCTGAG	TTCCCACTGG	1620
	CTAGCATCGG	CAAAAAGATG	GCTAATGCAC	AGAGGACAAC	AGTAGATTTG	AACGAGGCTC	1680
	CTTTCAAGAT	AAAAGAGGAG	CACTTGAATC	GGCTTAGAGC	ACTCTCTAGA	ACTGTAGAAC	1740
35	TTGGAAAACG	CTTCTTTCCA	CGTTGTTTCA	AAGTTCTAAA	TAAGATCATG	GATGCTGATG	1800
	ACTTGTCTGA	GATAGCTTAC	ATGGGGAATG	ATACGGCAGA	AGAGCGTCAA	CTGAAGAAGC	1860
	AAAGGTACAT	GGAACTTCAA	GAAATTCTGA	CTAAAGCATT	CACTGAGGAT	AAAGAAGAAT	1920
	ATGATAAGAC	TAACAACATC	TCCTCATCTT	GTTCTCTTAC	ATCTAAGGGA	GATAGATAAGC	1980
	CCAATAAGCT	CCCTTTTAGG	AAATAGGTAA	TTGTATTAGG	ATATATGAGG	AAGAAGAGGA	2040
40	TTTTCTTGTA	ACATAGCACT	CTTTCCTTTC	ATCATTTGAT	ATGTCAACAT	ACATACAACA	2100
	GCTGTACCAT	AAACTTGTAT	TGTTGCACTT	ACAACCTTGA	AGAACAGAAT	TTATTTGAAA	2160
	AAAAAAAAAA	AA					2172

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 588 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

	Met	Asp	Asn	Ser	Arg	Thr	Ala	Phe	Ser	Asp	Ser	Asn	Asp	Ile	Ser	Gly
	1				5				10					15		
5	Ser	Ser	Ser	Ile	Cys	Cys	Ile	Gly	Gly	Gly	Met	Thr	Glu	Phe	Phe	Ser
				20					25					30		
	Pro	Glu	Thr	Ser	Pro	Ala	Glu	Ile	Thr	Ser	Leu	Lys	Arg	Leu	Ser	Glu
			35					40					45			
	Thr	Leu	Glu	Ser	Ile	Phe	Asp	Ala	Ser	Leu	Pro	Glu	Phe	Asp	Tyr	Phe
		50					55					60				
10	Ala	Asp	Ala	Lys	Leu	Val	Val	Ser	Gly	Pro	Cys	Lys	Glu	Ile	Pro	Val
	65					70				75					80	
	His	Arg	Cys	Ile	Leu	Ser	Ala	Arg	Ser	Pro	Phe	Phe	Lys	Asn	Leu	Phe
				85					90					95		
15	Cys	Gly	Lys	Lys	Glu	Lys	Asn	Ser	Ser	Lys	Val	Glu	Leu	Lys	Glu	Val
				100					105					110		
	Met	Lys	Glu	His	Glu	Val	Ser	Tyr	Asp	Ala	Val	Met	Ser	Val	Leu	Ala
		115						120				125				
	Tyr	Leu	Tyr	Ser	Gly	Lys	Val	Arg	Pro	Ser	Pro	Lys	Asp	Val	Cys	Val
		130				135						140				
20	Cys	Val	Asp	Asn	Asp	Cys	Ser	His	Val	Ala	Cys	Arg	Pro	Ala	Val	Ala
	145					150				155					160	
	Phe	Leu	Val	Glu	Val	Leu	Tyr	Thr	Ser	Phe	Thr	Phe	Gln	Ile	Ser	Glu
				165						170					175	
25	Leu	Val	Asp	Lys	Phe	Gln	Arg	His	Leu	Leu	Asp	Ile	Leu	Asp	Lys	Thr
			180						185				190			
	Ala	Ala	Asp	Asp	Val	Met	Met	Val	Leu	Ser	Val	Ala	Asn	Ile	Cys	Gly
		195					200					205				
	Lys	Ala	Cys	Glu	Arg	Leu	Leu	Ser	Ser	Cys	Ile	Glu	Ile	Ile	Val	Lys
		210				215				220						
30	Ser	Asn	Val	Asp	Ile	Ile	Thr	Leu	Asp	Lys	Ala	Leu	Pro	His	Asp	Ile
	225					230				235					240	
	Val	Lys	Gln	Ile	Thr	Asp	Ser	Arg	Ala	Glu	Leu	Gly	Leu	Gln	Gly	Pro
				245						250				255		
35	Glu	Ser	Asn	Gly	Phe	Pro	Asp	Lys	His	Val	Lys	Arg	Ile	His	Arg	Ala
			260					265					270			
	Leu	Asp	Ser	Asp	Asp	Val	Glu	Leu	Leu	Gln	Met	Leu	Leu	Arg	Glu	Gly
		275					280					285				
	His	Thr	Thr	Leu	Asp	Asp	Ala	Tyr	Ala	Leu	His	Tyr	Ala	Val	Ala	Tyr
		290				295						300				
40	Cys	Asp	Ala	Lys	Thr	Thr	Ala	Glu	Leu	Leu	Asp	Leu	Ala	Leu	Ala	Asp
	305					310				315					320	
	Ile	Asn	His	Gln	Asn	Ser	Arg	Gly	Tyr	Thr	Val	Leu	His	Val	Ala	Ala
				325						330				335		
45	Met	Arg	Lys	Glu	Pro	Lys	Ile	Val	Val	Ser	Leu	Leu	Thr	Lys	Gly	Ala
			340						345				350			
	Arg	Pro	Ser	Asp	Leu	Thr	Ser	Asp	Gly	Arg	Lys	Ala	Leu	Gln	Ile	Ala
		355					360					365				
	Lys	Arg	Leu	Thr	Arg	Leu	Val	Asp	Phe	Ser	Lys	Ser	Pro	Glu	Glu	Gly
		370				375					380					
50	Lys	Ser	Ala	Ser	Asn	Asp	Arg	Leu	Cys	Ile	Glu	Ile	Leu	Glu	Gln	Ala
	385					390					395				400	
	Glu	Arg	Arg	Asp	Pro	Leu	Leu	Gly	Glu	Ala	Ser	Val	Ser	Leu	Ala	Met
				405					410				415			
	Ala	Gly	Asp	Asp	Leu	Arg	Met	Lys	Leu	Leu	Tyr	Leu	Glu	Asn	Arg	Val

420 425 430
 Gly Leu Ala Lys Leu Leu Phe Pro Met Glu Ala Lys Val Ala Met Asp
 435 440 445
 5 Ile Ala Gln Val Asp Gly Thr Ser Glu Phe Pro Leu Ala Ser Ile Gly
 450 455 460
 Lys Lys Met Ala Asn Ala Gln Arg Thr Thr Val Asp Leu Asn Glu Ala
 465 470 475 480
 Pro Phe Lys Ile Lys Glu Glu His Leu Asn Arg Leu Arg Ala Leu Ser
 485 490 495
 10 Arg Thr Val Glu Leu Gly Lys Arg Phe Phe Pro Arg Cys Ser Glu Val
 500 505 510
 Leu Asn Lys Ile Met Asp Ala Asp Asp Leu Ser Glu Ile Ala Tyr Met
 515 520 525
 15 Gly Asn Asp Thr Ala Glu Glu Arg Gln Leu Lys Lys Gln Arg Tyr Met
 530 535 540
 Glu Leu Gln Glu Ile Leu Thr Lys Ala Phe Thr Glu Asp Lys Glu Glu
 545 550 555 560
 Tyr Asp Lys Thr Asn Asn Ile Ser Ser Ser Cys Ser Ser Thr Ser Lys
 565 570 575
 20 Gly Val Asp Lys Pro Asn Lys Leu Pro Phe Arg Lys
 580 585

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GTGACAGACT TGCTCCTACT G

21

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CAGTGTGTAT CAAAGCACCA

20

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

5 TTCTCCAGAC CACATGATTA T 21

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TGAAGCTAAT ATGCACAGGA G 21

15 (2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GTAGGTGCTC TTGTTCTTCC C 21

(2) INFORMATION FOR SEQ ID NO:20:

25 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE:DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CACATAATTC CCACGAGGAT C 21

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 17 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Lys Gly Thr Cys Glu Phe Ile Val Thr Ser Leu Glu Pro Asp Arg
1 5 10 15
Leu

10 (2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
15 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Arg Arg Lys Glu Leu Gly Leu Glu Val Pro Lys Val Lys Lys
1 5 10

20 (2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Lys Lys Gln Arg Tyr Met Glu Ile Gln Glu Thr Leu Lys Lys
1 5 10

30 (2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
35 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AARGARGAYC AYACNAA

17

(2) INFORMATION FOR SEQ ID NO:25:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 17 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:
 TAYGTYAAYG TNAARAC 17
 (2) INFORMATION FOR SEQ ID NO:26:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 17 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:
 GCCATNGTNG CYTGYTT 17
 (2) INFORMATION FOR SEQ ID NO:27:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 17 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:
 AARGTNAARA ARCAYGT 17
 (2) INFORMATION FOR SEQ ID NO:28:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE DNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:
 RAAAYTCRCAN GTNCCYTTCA T 21